



wwPDB X-ray Structure Validation Summary Report ⓘ

May 18, 2020 – 02:09 am BST

PDB ID : 6JFB
Title : Crystal structure of human pyruvate kinase M2 isoform
Authors : Chen, T.J.; Wang, W.C.
Deposited on : 2019-02-08
Resolution : 2.12 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.11

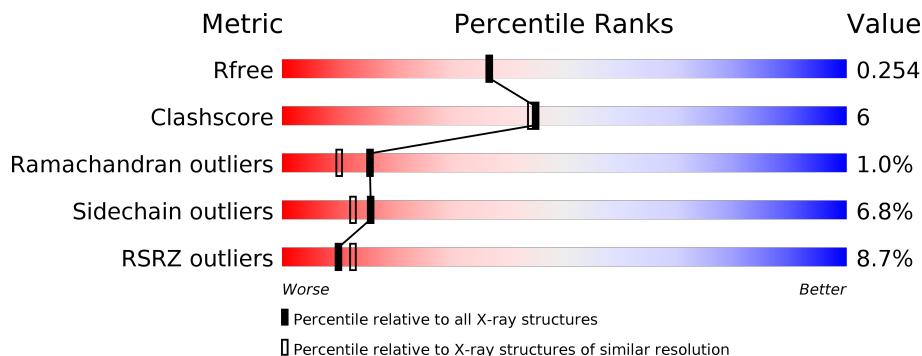
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.12 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	6241 (2.14-2.10)
Clashscore	141614	6778 (2.14-2.10)
Ramachandran outliers	138981	6705 (2.14-2.10)
Sidechain outliers	138945	6706 (2.14-2.10)
RSRZ outliers	127900	6112 (2.14-2.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	552	 9% 78% 15% • 5%
1	B	552	 7% 78% 14% •• 6%
1	C	552	 7% 77% 14% • 5%
1	D	552	 11% 77% 14% • 7%

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 16478 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Pyruvate kinase PKM Isoform M2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	522	3997	2515	708	749	25	0	0	0
1	B	521	3994	2513	707	749	25	0	0	0
1	C	523	4006	2520	709	752	25	0	0	0
1	D	516	3950	2485	701	740	24	0	0	0

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	expression tag	UNP P14618
A	-19	GLY	-	expression tag	UNP P14618
A	-18	SER	-	expression tag	UNP P14618
A	-17	SER	-	expression tag	UNP P14618
A	-16	HIS	-	expression tag	UNP P14618
A	-15	HIS	-	expression tag	UNP P14618
A	-14	HIS	-	expression tag	UNP P14618
A	-13	HIS	-	expression tag	UNP P14618
A	-12	HIS	-	expression tag	UNP P14618
A	-11	HIS	-	expression tag	UNP P14618
A	-10	SER	-	expression tag	UNP P14618
A	-9	SER	-	expression tag	UNP P14618
A	-8	GLY	-	expression tag	UNP P14618
A	-7	LEU	-	expression tag	UNP P14618
A	-6	VAL	-	expression tag	UNP P14618
A	-5	PRO	-	expression tag	UNP P14618
A	-4	ARG	-	expression tag	UNP P14618
A	-3	GLY	-	expression tag	UNP P14618
A	-2	SER	-	expression tag	UNP P14618
A	-1	HIS	-	expression tag	UNP P14618
A	0	MET	-	expression tag	UNP P14618

Continued on next page...

Continued from previous page...

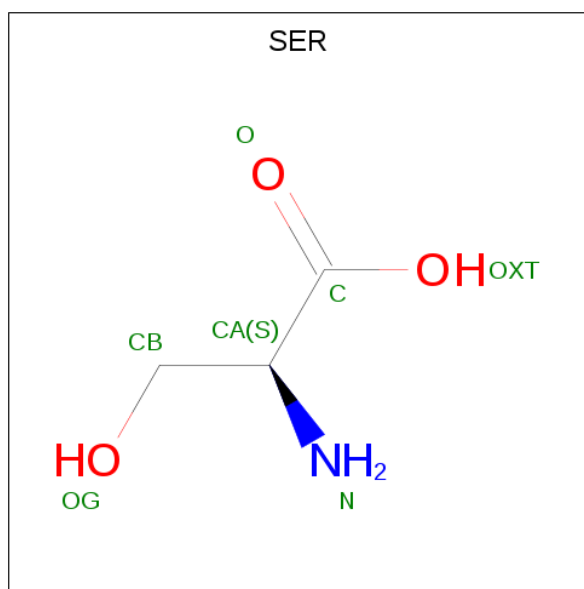
Chain	Residue	Modelled	Actual	Comment	Reference
B	-20	MET	-	expression tag	UNP P14618
B	-19	GLY	-	expression tag	UNP P14618
B	-18	SER	-	expression tag	UNP P14618
B	-17	SER	-	expression tag	UNP P14618
B	-16	HIS	-	expression tag	UNP P14618
B	-15	HIS	-	expression tag	UNP P14618
B	-14	HIS	-	expression tag	UNP P14618
B	-13	HIS	-	expression tag	UNP P14618
B	-12	HIS	-	expression tag	UNP P14618
B	-11	HIS	-	expression tag	UNP P14618
B	-10	SER	-	expression tag	UNP P14618
B	-9	SER	-	expression tag	UNP P14618
B	-8	GLY	-	expression tag	UNP P14618
B	-7	LEU	-	expression tag	UNP P14618
B	-6	VAL	-	expression tag	UNP P14618
B	-5	PRO	-	expression tag	UNP P14618
B	-4	ARG	-	expression tag	UNP P14618
B	-3	GLY	-	expression tag	UNP P14618
B	-2	SER	-	expression tag	UNP P14618
B	-1	HIS	-	expression tag	UNP P14618
B	0	MET	-	expression tag	UNP P14618
C	-20	MET	-	expression tag	UNP P14618
C	-19	GLY	-	expression tag	UNP P14618
C	-18	SER	-	expression tag	UNP P14618
C	-17	SER	-	expression tag	UNP P14618
C	-16	HIS	-	expression tag	UNP P14618
C	-15	HIS	-	expression tag	UNP P14618
C	-14	HIS	-	expression tag	UNP P14618
C	-13	HIS	-	expression tag	UNP P14618
C	-12	HIS	-	expression tag	UNP P14618
C	-11	HIS	-	expression tag	UNP P14618
C	-10	SER	-	expression tag	UNP P14618
C	-9	SER	-	expression tag	UNP P14618
C	-8	GLY	-	expression tag	UNP P14618
C	-7	LEU	-	expression tag	UNP P14618
C	-6	VAL	-	expression tag	UNP P14618
C	-5	PRO	-	expression tag	UNP P14618
C	-4	ARG	-	expression tag	UNP P14618
C	-3	GLY	-	expression tag	UNP P14618
C	-2	SER	-	expression tag	UNP P14618
C	-1	HIS	-	expression tag	UNP P14618
C	0	MET	-	expression tag	UNP P14618

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-20	MET	-	expression tag	UNP P14618
D	-19	GLY	-	expression tag	UNP P14618
D	-18	SER	-	expression tag	UNP P14618
D	-17	SER	-	expression tag	UNP P14618
D	-16	HIS	-	expression tag	UNP P14618
D	-15	HIS	-	expression tag	UNP P14618
D	-14	HIS	-	expression tag	UNP P14618
D	-13	HIS	-	expression tag	UNP P14618
D	-12	HIS	-	expression tag	UNP P14618
D	-11	HIS	-	expression tag	UNP P14618
D	-10	SER	-	expression tag	UNP P14618
D	-9	SER	-	expression tag	UNP P14618
D	-8	GLY	-	expression tag	UNP P14618
D	-7	LEU	-	expression tag	UNP P14618
D	-6	VAL	-	expression tag	UNP P14618
D	-5	PRO	-	expression tag	UNP P14618
D	-4	ARG	-	expression tag	UNP P14618
D	-3	GLY	-	expression tag	UNP P14618
D	-2	SER	-	expression tag	UNP P14618
D	-1	HIS	-	expression tag	UNP P14618
D	0	MET	-	expression tag	UNP P14618

- Molecule 2 is SERINE (three-letter code: SER) (formula: C₃H₇NO₃).



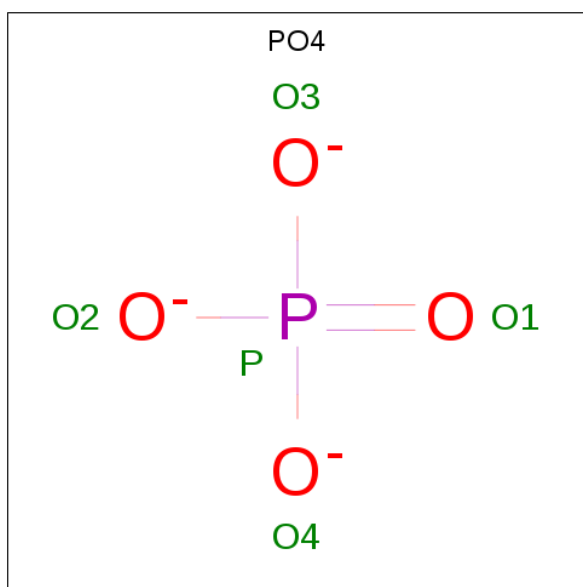
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
			Total	C	N			O
2	A	1	7	3	1	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	N	O	0	0
			7	3	1	3		
2	C	1	Total	C	N	O	0	0
			7	3	1	3		
2	D	1	Total	C	N	O	0	0
			7	3	1	3		

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	O P	0	0
			5	4 1		
3	B	1	Total	O P	0	0
			5	4 1		
3	C	1	Total	O P	0	0
			5	4 1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	88	Total	O	0	0
			88	88		
4	B	156	Total	O	0	0
			156	156		
4	C	134	Total	O	0	0
			134	134		

Continued on next page...

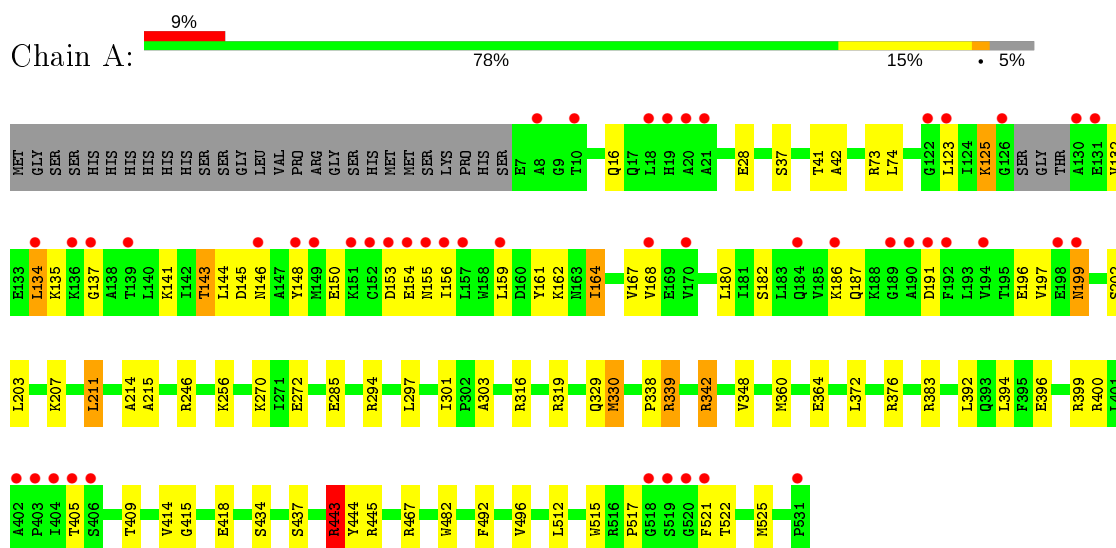
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	110	Total 110	O 110	0	0

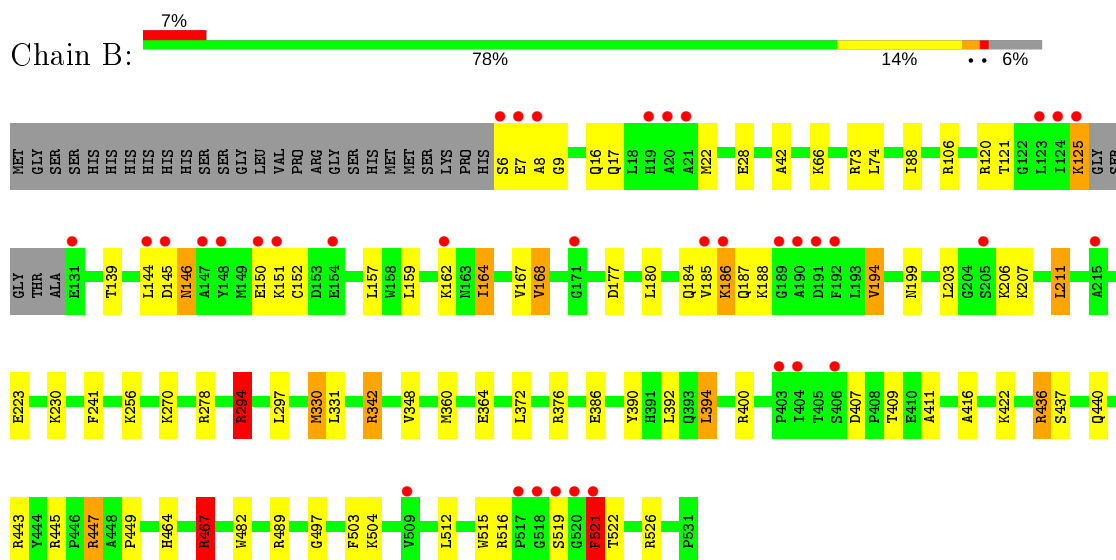
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

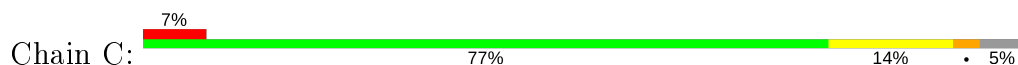
● Molecule 1: Pyruvate kinase PKM Isoform M2

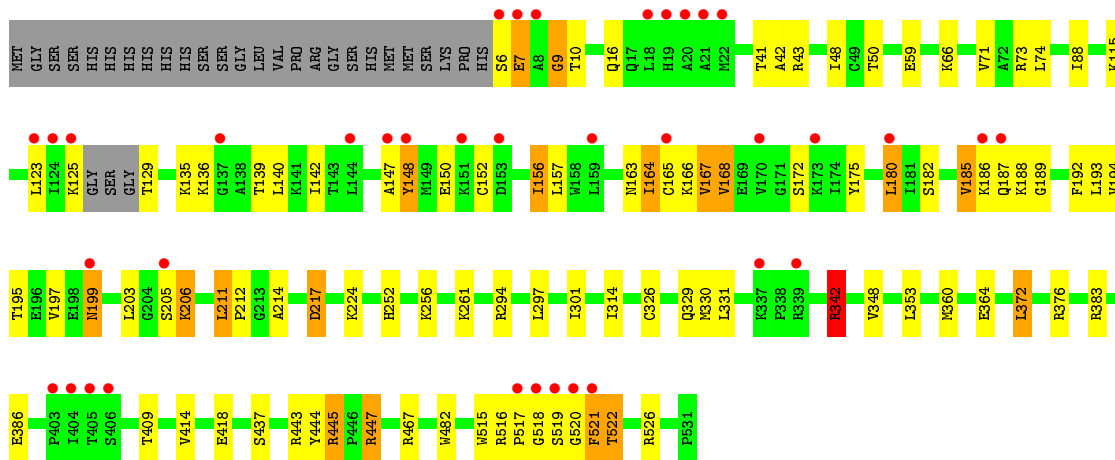


● Molecule 1: Pyruvate kinase PKM Isoform M2

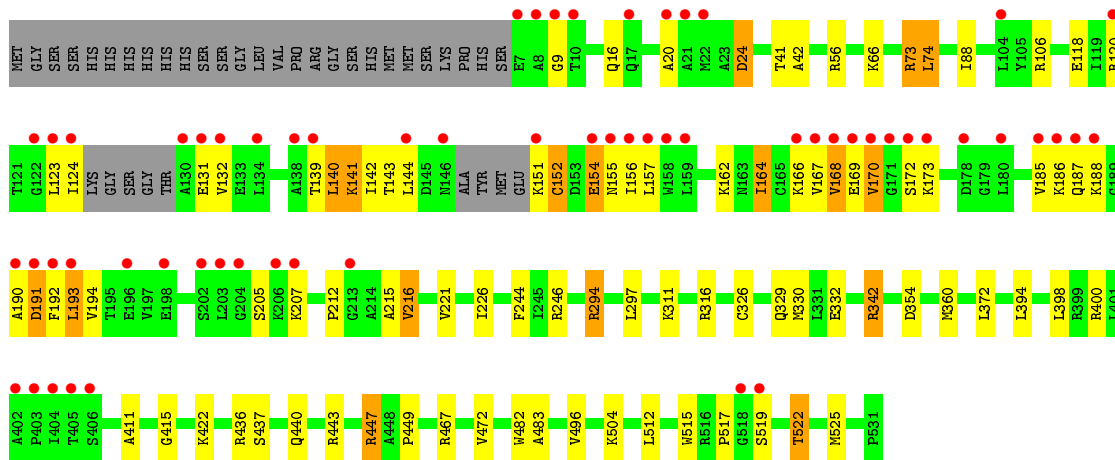
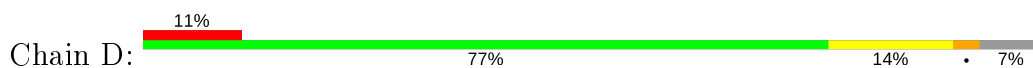


● Molecule 1: Pyruvate kinase PKM Isoform M2





• Molecule 1: Pyruvate kinase PKM Isoform M2



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	76.76Å 152.57Å 94.41Å 90.00° 97.88° 90.00°	Depositor
Resolution (Å)	29.96 – 2.12 29.94 – 2.12	Depositor EDS
% Data completeness (in resolution range)	99.5 (29.96-2.12) 99.6 (29.94-2.12)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.48 (at 2.12Å)	Xtrriage
Refinement program	REFMAC 5.8.0230	Depositor
R, R_{free}	0.192 , 0.251 0.199 , 0.254	Depositor DCC
R_{free} test set	6128 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	24.0	Xtrriage
Anisotropy	0.070	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 48.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	16478	wwPDB-VP
Average B, all atoms (Å ²)	34.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PO4

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.58	1/4061 (0.0%)	0.72	1/5483 (0.0%)
1	B	0.60	0/4058	0.75	1/5479 (0.0%)
1	C	0.61	0/4070	0.74	1/5496 (0.0%)
1	D	0.59	0/4012	0.72	1/5417 (0.0%)
All	All	0.59	1/16201 (0.0%)	0.73	4/21875 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	9
1	B	0	6
1	C	0	5
1	D	0	6
All	All	0	26

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	364	GLU	CD-OE1	-5.05	1.20	1.25

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	364	GLU	CB-CA-C	-6.13	98.15	110.40
1	D	106	ARG	NE-CZ-NH1	-5.50	117.55	120.30
1	B	467	ARG	NE-CZ-NH1	-5.24	117.68	120.30
1	C	445	ARG	NE-CZ-NH2	-5.05	117.78	120.30

There are no chirality outliers.

5 of 26 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	246	ARG	Sidechain
1	A	316	ARG	Sidechain
1	A	339	ARG	Sidechain
1	A	376	ARG	Sidechain
1	A	383	ARG	Sidechain

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3997	0	4079	52	0
1	B	3994	0	4076	61	1
1	C	4006	0	4088	55	0
1	D	3950	0	4033	57	1
2	A	7	0	4	0	0
2	B	7	0	4	1	0
2	C	7	0	4	0	0
2	D	7	0	4	0	0
3	A	5	0	0	1	0
3	B	5	0	0	0	0
3	C	5	0	0	0	0
4	A	88	0	0	1	0
4	B	156	0	0	0	0
4	C	134	0	0	4	0
4	D	110	0	0	2	0
All	All	16478	0	16292	209	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

The worst 5 of 209 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:372:LEU:HD11	1:B:376:ARG:CZ	1.85	1.07

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:339:ARG:NH1	1:B:199:ASN:HD21	1.56	1.02
1:B:436:ARG:HH21	1:B:443:ARG:HH12	1.12	0.97
1:B:436:ARG:HH21	1:B:443:ARG:NH1	1.67	0.93
1:A:339:ARG:HH11	1:B:199:ASN:HD21	0.91	0.88

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:278:ARG:NH2	1:D:166:LYS:O[1_556]	2.12	0.08

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	518/552 (94%)	493 (95%)	20 (4%)	5 (1%)	15 10
1	B	517/552 (94%)	495 (96%)	16 (3%)	6 (1%)	13 8
1	C	519/552 (94%)	494 (95%)	20 (4%)	5 (1%)	15 10
1	D	510/552 (92%)	487 (96%)	18 (4%)	5 (1%)	15 10
All	All	2064/2208 (94%)	1969 (95%)	74 (4%)	21 (1%)	15 10

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	125	LYS
1	A	154	GLU
1	C	148	TYR
1	C	519	SER
1	D	140	LEU

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	428/454 (94%)	397 (93%)	31 (7%)	14	11
1	B	429/454 (94%)	400 (93%)	29 (7%)	16	12
1	C	430/454 (95%)	398 (93%)	32 (7%)	13	10
1	D	424/454 (93%)	399 (94%)	25 (6%)	19	16
All	All	1711/1816 (94%)	1594 (93%)	117 (7%)	16	12

5 of 117 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	330	MET
1	C	156	ILE
1	D	294	ARG
1	B	392	LEU
1	B	521	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	17	GLN
1	C	187	GLN
1	D	29	HIS
1	C	16	GLN
1	D	155	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

7 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
3	PO4	C	602	-	4,4,4	0.68	0	6,6,6	1.02	0
3	PO4	B	602	-	4,4,4	0.60	0	6,6,6	1.44	1 (16%)
3	PO4	A	602	-	4,4,4	0.97	0	6,6,6	1.14	0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
3	B	602	PO4	O4-P-O1	-2.17	102.95	110.89

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	602	PO4	1	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	522/552 (94%)	0.42	47 (9%) 9 12	14, 30, 81, 116	0
1	B	521/552 (94%)	0.25	36 (6%) 16 20	11, 24, 64, 95	0
1	C	523/552 (94%)	0.27	37 (7%) 16 19	12, 26, 63, 92	0
1	D	516/552 (93%)	0.50	61 (11%) 4 5	14, 29, 90, 131	0
All	All	2082/2208 (94%)	0.36	181 (8%) 10 12	11, 27, 76, 131	0

The worst 5 of 181 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	519	SER	10.2
1	D	130	ALA	7.8
1	D	171	GLY	7.2
1	B	520	GLY	7.0
1	B	404	ILE	6.5

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	PO4	A	602	5/5	0.80	0.24	35,44,59,65	0
2	SER	D	601	7/7	0.88	0.17	27,29,30,31	0
2	SER	C	601	7/7	0.92	0.15	21,23,24,26	0
3	PO4	C	602	5/5	0.93	0.15	28,35,39,40	0
2	SER	A	601	7/7	0.94	0.16	23,24,28,33	0
2	SER	B	601	7/7	0.95	0.21	16,19,22,25	0
3	PO4	B	602	5/5	0.96	0.15	33,40,43,45	0

6.5 Other polymers [i](#)

There are no such residues in this entry.